# On Monitoring of End-to-End Packet Reordering over the Internet[*]

Bin Ye[1] Anura P. Jayasumana[1] Nischal M. Piratla[2]
1Computer Networking Research laboratory, Colorado State University, Fort Collins, CO 80523, USA
2 Deutsche Telekom Laboratories, Ernst-Reuter-Platz 7, D-10587 Berlin, Germany
Email: {Bin.Ye, Anura.Jayasumana}@colostate.edu, Nischal.Piratla@telekom.de

## Abstract

*End-to-end reordering of packets on Internet is investigated. Packet streams transferred over the Internet are used to analyze the long-term and short-term trends in reordering. Reorder Density (RD) is used to capture comprehensively and concisely the nature of reordering present in a stream. Simpler metrics are derived from RD for monitoring of long-term reordering trends. Reorder Entropy characterizes the reordering in a stream using a single value, reflecting the fraction of packets displaced as well as the magnitude displacements. The mean displacement of packets, defined with respect to late packets, early packets or all the packets, together with the percentage of packets that are displaced can also be used for this purpose. The pros and cons of each of these for monitoring long-term trends in reordering are addressed. The measurements presented indicate that for some end-to-end links the packet reordering shows clear daily and weekly trends.*

## 1. Introduction

The Internet Protocol (IP) provides a best-effort packet delivery service, and as such there is no guarantee on the delivery of packets or the order in which packets are delivered. Reordering of packets in a stream is not uncommon in the Internet, and is not a sign of pathological behavior of the Internet. Increase in parallelism required in nodes to handle high link speeds and large routing tables can cause reordering of packets within a routing node. The major cause of reordering is the parallelism, both within the nodes (switches and routers) and among the links[1,3,13]. Load balancing policies within networks introduce switch-level parallelism. Route flapping, where routes between autonomous systems (AS) change frequently, also causes different packets of the same stream to take different paths possibly leading to out-of-order delivery. In wireless communication, the order of delivery of packets is further aggravated by hand-offs.

With UDP the packets are delivered to the application in the same order that they arrive at the IP layer, thus forcing the application to deal with reordering unless the application itself is not sensitive to the order of packets. For delay sensitive applications based on UDP, such as videoconferencing, an out-of-order packet that arrives after the elapse of playback time is as good as lost. Even though TCP provides in-order delivery of the stream to the application, reordering at the IP level can still significantly impact the performance of TCP based applications [4, 5,8,13]. For example, reordering of three or more packet positions within a flow of a TCP connection may cause fast retransmission and fast recovery multiple times, resulting in a reduced congestion window and consequently a drop in throughput for the application [4]. Despite efforts of network engineers and operators to limit reordering, it continues to exist due to the intrinsic dynamics and parallelism of traffic in packet switched networks [2,12]. As the network speeds increase the amount of reordering can also be expected to increase.

End-to-end measurements of packet streams provide a valuable perspective as they indicate the conditions that an application associated flow experiences over the Internet paths [9]. Since the validity of active measurements associated with the study of end-to-end reordering characteristics depends heavily on the sampling process, a passive measurement approach is used in this study. Measurements are repeated at different times of the day as well as on different days of the week. The goal of this research is to characterize reordering observed in the Internet. Specifically, it investigates whether there are daily or weekly trends associated with packet reordering. To achieve this, we first define simple metrics to easily identify the end-to-end packet reordering trends. Next we provide an empirical examination of the end-to-end packet reordering measurements over long periods by using the proposed metrics.

---

[*] This research was supported in part by NSF ITR Grant No.0121546, and Ixia University Partners Program.

Section 2 briefly reviews Reorder Density (RD), the metric used for measurements, along with its basic properties. Section 3 elaborates on the utility of RD for classification and introduces several simple derived metrics for long-term monitoring of reordering. Section 4, presents the experimental methodology.. In Section 5, we present the measurements, compare the proposed metrics, and examine end-to-end packet reordering trends. Section 6 presents the conclusions.

## 2. Reorder Density

Metrics proposed for measurement of reordering, include Reorder Buffer-occupancy Density (RBD), Reorder Extent, n-Reordering and Reorder Density [14]. Some are mere measures that vary with packet reordering, and fail to provide insight into the nature of reordering. Reorder Buffer occupancy Density (RBD) [1] evaluates the reordering from the point of view of resources needed for recovery from reordering. Reorder Density RD comprehensively captures packet reordering, and has been demonstrated to have several useful properties [14]. Its attributes include informativeness, low evaluation complexity, and robustness. It is unique among the metrics in its extensibility to cascades of networks, a feature that is essential for monitoring complex networks. However due to its comprehensiveness of expression, i.e., non-atomic nature, simpler "derived" metrics would be more convenient for monitoring end-to-end links for reordering over long time periods. Next we provide a brief description of RD [13].

Consider a sequence of packets $(1,2,3,4,...N)$ transmitted over a network. A 'receive index' $(1,2,3,4,...N)$ is assigned to each packet as it arrives at the destination. Lost and duplicate packets are not assigned a receive index. Let the receive index assigned to packet m be $(m + d_m)$. In the absence of reordering, the sequence number of the packet and the receive index are same, i.e. $d_m = 0$ for each packet. When $d_m \neq 0$ we say that a reorder event has occurred, and denote this event by $r(m, d_m)$. If $d_m > 0$, the packet is considered late, and if $d_m < 0$, the packet is early.

Let $S[k]$ denote the set of reorder events with displacement equal to $k$, i.e., $S[k] = \{r(m, d_m) | d_m = k\}$. $RD[k]$ is defined as the discrete density of the frequency of packets with respect to their displacements normalized to the total number of non-duplicate received packets, $N^{'}$ Thus,

$$RD[k] = \frac{|S[k]|}{N^{'}} \qquad (1)$$

RD provides the use of a threshold, $D_T$, beyond which an early or a late packet is deemed to be lost. Thus, in the above expression, $-D_T \leq k \leq D_T$. A sequence of packets (1, 2,…N) transmitted over a network is called one observation. By maintaining a threshold, $D_T$ , on displacement of any packet, and as well as an early arrival buffer, RD remains insensitive to lost and duplicate packets, while maintaining the real-time evaluation of reordering [6]. Several properties of RD make it a useful and a comprehensive metric [6,13,14]. RD[*i*] satisfies the following:

$$RD[i] \geq 0, \forall i \qquad (2)$$

$$\sum_i RD[i] = 1 \qquad (3)$$

$$\sum_i (i \times RD[i]) = 0 \qquad (4)$$

While it is possible to monitor the shape of RD, to monitor reordering in a network over long term, one may be interested in a simpler metric. The use of RD vs. an appropriate, simple derived metric for monitoring long term trends could be considered analogous to monitoring the quality of a communication link using spectrum analyzer vs. the signal-to-noise ratio measurements. While the former provides detailed information on noise spectrum at a given instant, for long-term quality measurement, the SNR is more convenient and often sufficient.

## 3. Simple Reorder Metrics

In this section, we present several simple, derived metrics to monitor packet reordering in a network or an end-to-end connection over a long time period. Percentage of late packets ($P_L$), Mean displacement of packets ($M_D$), means displacement of late packets ($M_L$), and the reorder entropy ($E_R$). These metrics are formally defined next and their properties described.

### *a) Percentage of Late Packets ($P_L$)*

To capture the lateness of packets from their original positions the percentage of late packets is defined as the percentage of packets that exhibit lateness with respect to their expected position, as given by the receive index.

$$P_L = \sum_{i=+1}^{i=D_T} RD[i] \qquad (5)$$

$P_L$ = 0 corresponds to the case where all the packets are in order. Ref.[13] illustrates why a late-packet has to be associated with one or more early packets and an early packet has to be associated with one or more late packets. Thus, for a sequence with packet reordering, $P_L > 0$. Similarly, percentage of earliness ($P_E$) may be defined, and the percentage of earliness and percentage of lateness need not be equal.

### *b) Mean Displacement of Packets ($M_D$)*

Packet reordering is associated with two types of events, lateness events and earliness events. In a lateness event, the corresponding displacement is always positive. And a negative displacement is mapped with the earliness event. When calculating the mean displacement of packets, if both late and early packets are included, from equation (4), the mean displacement is zero for all cases. Therefore, the mean displacement, when all packets are taken together, is not useful. On the other hand, one can consider the magnitude of displacement of packets, and divide it by the total number of packets to define a mean displacement $M_D$:

$$M_D = \left|\sum_{i=-D_T}^{i=+D_T}(|i|\times RD[i])\right| / \left|\sum_{i=-D_T}^{i=+D_T} RD[i]\right| \tag{6}$$

### *Mean displacement of late packets ($M_L$)*

RD[*i*] refers to the probability that a packet arrives *i* packets away from its expected position. Thus considering only the late packets, the mean displacement of late packets is given as:

$$M_L = \left[\sum_{i=1}^{i=+D_T}(i\times RD[i])\right] / \left[\sum_{i=1}^{i=D_T} RD[i]\right] \tag{7}$$

Similarly, the mean displacement for earliness is:

$$M_E = \left|\left[\sum_{i=-1}^{i=-D_T} i\times RD[i]\right] / \left[\sum_{i=-1}^{i=-D_T} RD[i]\right]\right| \tag{8}$$

Note here that we divide the total positive (negative) displacement by the total number of late (early) packets. Both $M_L$ and $M_E$ are always none negative values, and $M_L = M_D/2P_L$, $M_E = M_D/2P_E$.

### *c) Reorder Entropy ($E_R$)*

Entropy is a concept that is used to define the randomness or the disorder. Consider a set of n events, $S_1, \ldots, S_n$, with probabilities $p_1, \ldots, p_n$ adding up to 1. The expected information content of a probability distribution, called entropy, is derived by weighing the information values by their respective probabilities [15]:

$$E = \sum_{i=1}^{n} p_i \log_e(1/p_i) \tag{9}$$

This work exploits entropy as a convenient summary statistic for a distribution's tendency to be concentrated or dispersed. As RD is a discrete probability distribution, that of packet displacement (a form of disorder), we define reorder entropy as:

$$E_R = (-1)\times\sum_{i=-D_T}^{i=+D_T}(RD[i]\times\log_e RD[i]) \tag{10}$$

When no packet ordering is present, i.e., $RD[0]=1$, the reorder entropy is equal to zero. On the other hand, the packet sequence has the most variance, when packets are displaced uniformly with equal probabilities [15]. Thus an upper bound for entropy for a given threshold $D_T$ is obtained as:

$$E_R = \ln(2D_T+1) \tag{11}$$

.

## 4. Measurement Methodology

Our experiment uses a passive network monitoring technique. HTTP or FTP traffic generation is provoked with requests for downloads from an observation point, located at CNRL, Colorado. Five source web servers were chosen in Africa, Asia, Europe and America to attain a wide geographical distribution. The addresses associated with each site, hop count to each site from monitoring point, and the respective mean round trip time (RTT) and standard deviation of the RTT (SD) (in milliseconds) are listed below*:

- Net-1: ftp.chg.ru , RTT = 181.56, SD = 3.68, 13 Hops
- Net-2: ftp.debian.skynet.be, RTT = 135, SD = 2, 20 Hops
- Net-3: www.mara.org.za, RTT = 317.36, SD = 22.65, 17 Hops
- Net-4: informatics.nic.in, RTT = 334.5, SD = 2.5, 21 Hops
- Net-5: www.olympus.co.jp, RTT = 144.56, SD = 1.05, 13 Hops

* The readings listed were obtained from lamar.colostate.edu on 9:05am (MST), 28th April, 2006,from a combination of 'ping' and 'traceroute' outputs.

To guarantee that the connection between the source and destination are not short-lived, i.e., enough packets can be collected for analysis, only downloadable files with size larger than 2MB were chosen. Traces were gathered using 'tcpdump' and sequence numbers extracted. The first screening measurement was taken for 9 days in August, 2005. During the measurement period, the path from each web server to the CNRL host was probed using 'traceroute' utility periodically. We used 'wget' to initialize the packets transmissions between the source and destination every hour. Each collection captured at least 1,500 packets per stream. At the host side, the maximum download speed from each web site was 100 kBps. Scripts of RD available from [11] were used to parse the reorder density information. After analyzing the data obtained from the first set of screening experiments, a second set of measurement was carried out over a two-week period.

## 5. Results

### *a) Comparison of Simple Metrics:*

RD based measurements often can be characterized based on the shape of the function, e.g, *evenly distributed* vs. *unevenly distributed, narrow vs. wide and asymmetric vs. symmetric*. Fig. 1 illustrates an *evenly distributed* RD with $D_T$ = 25, observed on Net-5. It is found that all possible reordering events which fall within -25 to 25 have happened with no value above 0.04, and 32 out of 50's values above 0.02. No dominating displacement is present in this case. The entropy value (3.72) associated with this observation was the peak that day. A separate observation on Net-5 yielded an *unevenly distributed* RD, with RD [0] = 0.988, RD[-1]= 0.006, RD[1]=0.006, and all other RD components equal to zeros. For this case, entropy is equal to 0.073, $P_L$ = 0.006, and $M_L$= 1. The reordering events with -1 and +1 displacement were the only observed reorder events.

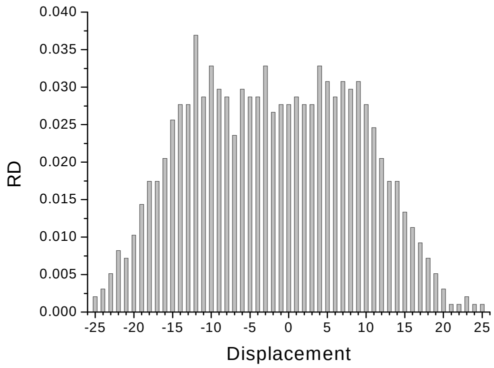


**Figure 1. Reorder Density on Net-5 observed at 9:52am August 26, 2005 ($D_T$ = 25)**

$P_L$ is the sum of lateness event distributions summarized in one measurement. If $P_L$ is low, then the chance of lateness events occurring is low. Same $P_L$ value could be associated with different entropy values. For example, points A and B in Fig. 2 have the same $P_L$ value (0.014), however, A's entropy is 0.40, three times that of B's. It can be seen from Fig. 3 that A has more displacement than B, which is an indication of higher amount of reordering. Thus, we can evaluate the magnitude of reordering with reorder entropy, for same $P_L$. In addition, entropy can capture the fluctuation of $M_D$, i.e., the weighted average displacement both for lateness and earliness events, as shown in Fig. 4.

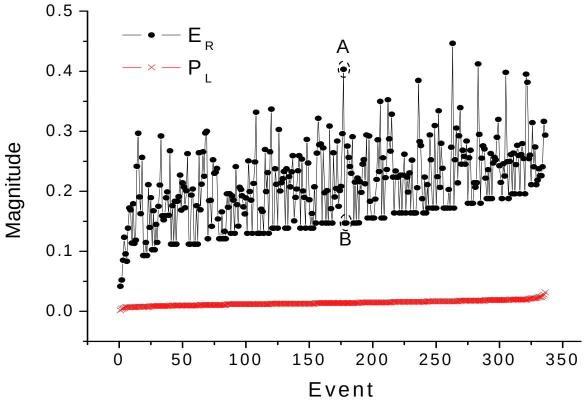


**Figure 2. $P_L$ and $E_R$ of Net – 5 (Sep 24- Oct 8, 2005); the points are sorted in ascending order of $P_L$.**

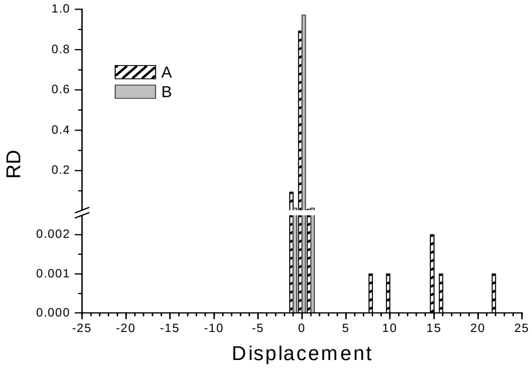


**Figure 3. RD of points A and B on figure 2.**

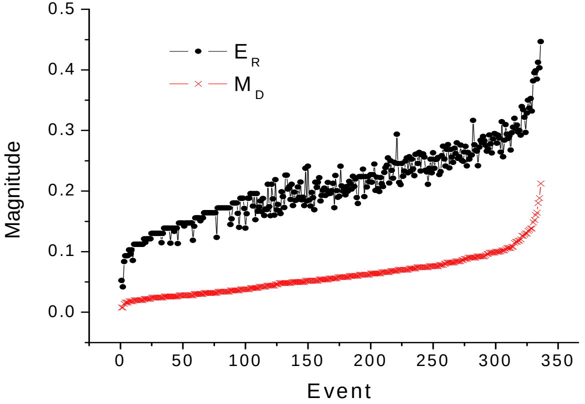


**Figure 4. $M_D$ and $E_R$ of Net – 5, (Sep 24 – Oct 8, 2005), points are sorted in ascending order of $M_D$ .**

To understand the relationship between the lateness and earliness events, $M_L$ and $M_E$ are examined in Fig. 5. In our experiments, no earliness event with the displacement' value > 2 was observed for Net-5. Over 90 percent of $M_E$ values observed were equal to 1. This type of event, a packet arriving early by one position, was common during the whole observation period. Fig. 5 shows that $M_L > M_E$ holds for 2/3 of measurement sets. This indicates that the reordering is caused by occasional packets getting delayed by more than one, thus causing a number of packets to arrive early by one. When $M_E = M_L = 1$, two adjacent packets switching their places is the dominant reorder event. However, the values of $M_E$ and $M_L$ do not convey the frequency with which it happens. For $M_E < M_L$ case, the receiving packets could be <1, 3, 4, 5, 6, 2, 7, 8, 9, 10>. In this scenario, a group of packets (3, 4, 5,6) arrive earlier than expected, as packet 2 is delayed by 4 positions. As a result, the sum of percentage of early packets $P_E$ is higher than $P_L$, the percentage of late packets. According to property 3 of RD, $M_E$ should be less than $M_L$. Note also that reorder entropy increases with the increase of $M_D$.

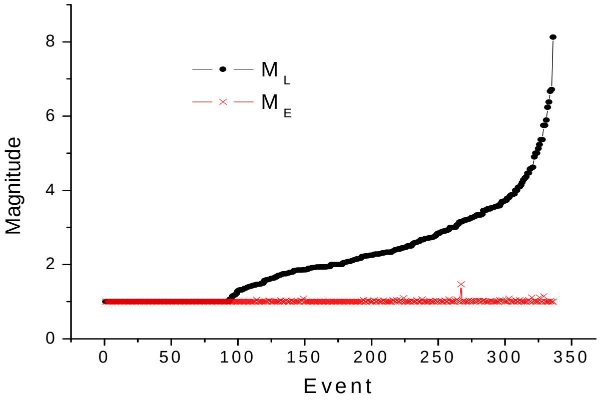


**Figure 5. $M_L$ and $M_E$ for Net-5 (Sep 24 – Oct 8, 2005), points are sorted in ascending order of $M_L$.**

### *b) Time Trends of Packet Reordering:*

In this section, $E_R$ is used to measure the end-to-end packer reordering of over a long period of time. Table 1 lists the mean of $E_R$, maximum $E_R$ and percentage of observations with $E_R > 0$ (i.e., reordering) during whole observation period for each network. One can see that packet reordering is prevalent in Net-1 and Net-5. Each packet sequence transmission over Net-5 throughout the observation period had reordering events. However, the maximum value of $E_R$ was not very high compared to other networks. The $E_R$ value of Net-5 was in the range [0.05, 0.45].

Another interesting result is that the end-to-end packet reordering on some paths does show similar trends over different days and different weeks. Figure 6 shows $E_R$, collected each hour, (Sep 24 -Oct 8, 2005) for Net-5. All clock times in this paper refer to the local time (Mountain Time) at which data was collected. Fig. 6a shows $E_R$ over two weekly periods. All the $E_R$ values observed were above zero, and the trend in the first week is very similar to that of the second. We extract the data for Monday, Tuesday and Wednesday, of each week and magnify them in Fig. 6b, c and d, respectively. The hourly entropy variation during these days in the two weeks shows similar time trends. Not only the absolute levels of entropy are fairly similar at a given time of day in the two weeks, but also increase and decrease of entropy happened during the hour on the particular day in the 2 weeks. However, not every network we measured show daily and weekly trends.

**Table 1: Statistical summary of packet reordering measurements (Sep 24 – Oct 8, 2005)**

| ***Net*** | ***Mean of $E_R$*** | ***Maximize $E_R$ Value*** | ***$E_R$ >0*** | |
|---|---|---|---|---|
| | | | ***Times*** | ***Percentage*** |
| 1 | 0.20 | 2.96 | 236 | 70.12 |
| 2 | 0.032 | 0.44 | 146 | 43.50 |
| 3 | 0.067 | 0.84 | 98 | 29.00 |
| 4 | 0.091 | 1.43 | 120 | 36.10 |
| 5 | 0.21 | 0.45 | 336 | 100 |

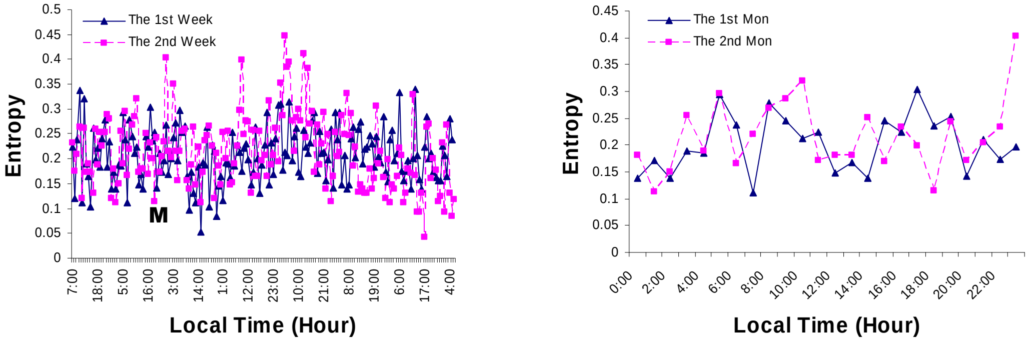
Note: The total number of measurement for each net is 336.

## 6. Summary

This research investigated the variation of reordering of packets over the Internet, and presented metrics for long-term reorder trend monitoring. While $P_L$, $M_D$, $M_L$ and $M_E$ and $E_R$ are all possible candidates for trend monitoring, reorder entropy is more useful and informative; it is sensitive not only to the change of amount of packet reordering, but also to the shape of the distribution of displacements. Use of reorder density and reorder entropy for reorder measurements is comparable to monitoring noise spectrum and Signal-to-Noise ratio respectively in monitoring quality of physical links. While the former provides detailed information that can be used to gain insight and diagnose problems, the latter is more convenient for long-term monitoring. End-to-end packet reordering measurements were performed over days and weeks. The results show that some paths exhibit daily and weekly reordering trends, while in others reordering is random without a clear trend. Our current research aims to investigate the causes of such behavior in detail. We also plan to do measurements over longer periods of time, and a variety of network links.

(A) Week Trend of Entropy

(B) Monday Trend



(C) Tuesday Trend

(D) Wednesday Trend

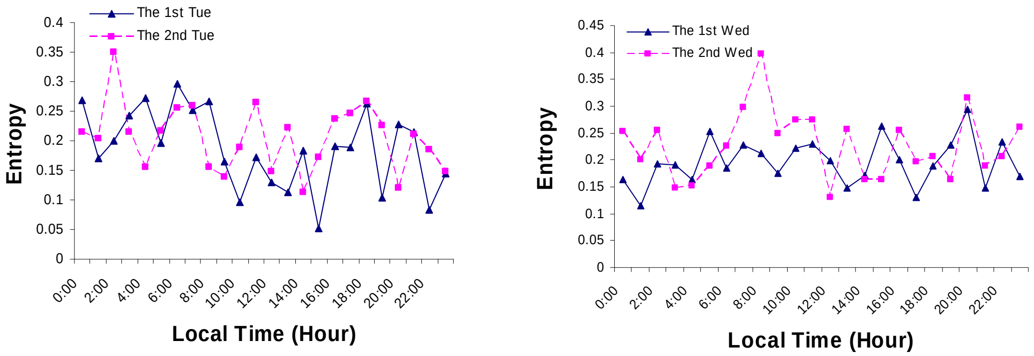


**Figure 6. Weekly and daily trends of end-to-end packet reordering on Net – 5 (Sep 24 – Oct 8, 2005)**

## References

[1] Banka, T., Bare, A. and Jayasumana, A., "Metrics for Degree of Reordering in Packet Sequences," *Proc. IEEE 27th Local Computer Networks Conf*, Nov. 2001, pp. 333-342.

[2] Bare, A. A., "Measurement and Analysis of Packet Reordering," Masters Thesis, Dep. Computer Science, Colorado State University, 2004.

[3] Bennett, J. C. R., Partridge, C. and Shectman, N., "Packet Reordering is Not Pathological Network Behavior," *Trans. on Networking IEEE/ACM*, Dec. 1999, pp.789-798.

[4] Blanton, E. and Allman, M., "On Making TCP More Robust to Packet Reordering", *ACM Computer Comm. Review*, 32(1), Jan. 2002, pp.20-30.

[5] Jaiswal, S., Iannaccone, G., Diot, C., Kurose, J. and Towsley, D., "Measurement and Classification of Out-of-sequence Packets in Tier-1 IP Backbone," *Proc. IEEE INFOCOM*, Mar. 2003, pp. 1199- 1209.

[6] Jayasumana, A., Piratla, N. M., Bare, A. A., Banka, T., Whitner R., and McCollom, J.,"Reorder Density Function - A Metric for Packet Reordering Measurement," IETF draft (work in progress), available at http://www.cnrl.colostate.edu /packet_reorder.html.

[7] Liu, H., "A Trace Driven Study of Packet Level Parallelism," *Proc. International Con. Communications (ICC'02)*, New York, NY, 2002, pp. 2191-2195.

[8] Loguinov, D. and Radha, H., "End-to-End Internet Video Traffic Dynamics: Statistical Study and Analysis," *Proc. IEEE INFOCOM*, Jun. 2002, pp.723-732.

[9] Luo X. and Chang R. K. C., "Novel Approaches to End-to-end Packet Reordering Measurement" *Proc. ACM/USENIX Conf. Internet Measurement*, Oct. 2005.

[10] Morton, A., Ciavattone, L., Ramachandran, G., Shalunov, S., and Perser, J., "Packet Reordering Metric for IPPM," IETF draft (work in progress).

[11] Perl scripts for RD, http://www.cnrl.colostate.edu /Reorder_perl_scripts.html. Last modified on Nov. 3, 2004.

[12] Piratla, N. M., "A Theoretical Foundation, Metrics and Modeling of Packet Reordering and Methodology of Delay Modeling using Inter-packet Gaps," *Ph.D. Dissertation*, Dept. Electrical and Computer Engineering, Colorado State University, Fall 2005.

[13] Piratla, N. M., Jayasumana, A. P., and Bare, A. A., "RD: A Formal, Comprehensive Metric for Packet Reordering," *Proceedings Fourth IFIP-TC6 Networking Conference (Networking 2005), LNCS 3462*, Ontario, May 2005, 78-89.

[14] Piratla, N. M. , Jayasumana A. P. ,and Bare A. A., "A Comparative Analysis of Packet Reordering Metrics," *Proc. IEEE/ACM 1st Int. Conf. Communication System Software and Middleware (COMSWARE 2006)*, New Delhi, Jan. 2006.

[15] Shannon C. E., "A Mathematical Theory of Communication", *Bell Sys. Tech. Journal*, vol. 27, 1948.